数字文献资源内容服务推荐研究*
——基于本体规则推理和语义相似度计算

刘 健[1] 毕 强[1] 刘庆旭[1] 王 福[1,2]
[1](吉林大学管理学院 长春 130022)
[2](内蒙古工业大学图书馆 呼和浩特 010051)

摘要: 【目的】解决传统数字文献资源内容服务推荐中无法充分挖掘资源语义信息等问题。【方法】通过设定本体推理规则对用户查询关键词进行语义扩展,提出一种新的语义相似度计算方法计算文献资源内容相似度。按照相似度大小对搜索结果进行排序,将排名较高的文献推荐给目标用户。【结果】实验结果证明,该方法能够较准确地计算语义相似度,并能够对用户需求进行有效推荐。【局限】缺少对数字资源的大规模采集,实验案例较少。【结论】该方法充分挖掘数字文献资源的语义信息并进行有效推荐,为数字资源内容服务推荐提供一种新思路。

关键词: 数字文献资源内容 服务推荐 本体推理 语义相似度

分类号: G250.7

1 引 言

以用户为中心,根据用户的个性化需求开展具有针对性和主动性的信息服务,是提高信息服务质量和信息资源使用效率的重要手段[1]。资源服务推荐是满足用户个性化价值追求的有效手段之一[2]。目前,基于内容的推荐[3]、基于协同过滤推荐[4-6]、以及基于情境的推荐[7-10]等资源服务推荐方法得到了较好的应用和推广。然而,上述推荐方法大多利用关键词词频作为计算依据,不能准确表达数字文献资源的语义信息[11],难以区分数字文献资源的品质和风格[12],导致数字文献资源的结构化程度较低,限制了资源的有效利用和共享[13],无法充分挖掘用户潜在的信息需求[14-15]。

本文将本体作为反映资源属性关系的模型,提出一种基于本体规则推理和语义相似度计算的数字文献资源推荐方法,以此解决推荐系统中资源的语义缺乏、结构化程度低等问题[16-17],并充分挖掘用户潜在需求,为数字资源内容服务推荐提供一种新思路。

2 相关研究

2.1 本体推荐

本体体现的是特定领域知识结构的概念体系,反映了特定领域的通用观点及其明确的概念和概念间关系的集合,侧重领域概念层次上的术语及术语关系的表达,为知识组织和共享提供精确控制[18]。基于本体的推荐方法可以有效解决推荐系统中语义缺乏、结构化程度低等问题[19],因而成为学者关注的热点。目前,主要有基于本体规则推理的资源推荐、基于本体语义相似度的资源推荐、基于本体语义描述的资源推荐等方式。基于本体规则推理的资源推荐是将语义网本体语言同推理机(如Jess、Pallet等)结合起来,通过设定、添加规则和进行逻辑推理,发掘隐含的语义关联关系,将关联度较高的资源推荐给目标用户[20-22]。基于本体语义相似度的资源推荐通过构建领域本体对多源信息进行整合,整合后的本体属性可以反映资源特征,再通过计算该本体中各个概念节点属性及本体网络结构的相

通讯作者:刘健,ORCID:0000-0001-8901-2814,E-mail:tomosliu9999@126.com。

*本文系国家自然科学基金项目"语义网络环境下数字图书馆资源多维度聚合与可视化展示研究"(项目编号:71273111)和"吉林大学高峰学科(群)建设项目"的研究成果之一。

似度从而得到相似资源，进而提高推荐质量[16, 23-25]。基于本体语义描述的资源推荐将本体引入到推荐系统中，使用OWL语言对用户和项目信息进行描述，然后与协同过滤、基于内容推荐等传统推荐模型相结合并计算用户偏好信息与项目信息的相似性。用户和项目具备了语义信息的同时，也提高了资源信息的结构化描述水平，并且可以有效提高推荐的召回率和准确率[26-28]。

然而基于本体规则推理的资源推荐仅通过设定推荐规则挖掘用户需求，但并未充分描述资源的语义信息；基于本体语义相似度资源推荐由于存在关键词和索引词之间的多重表达差异而无法表达用户的真实需求；基于本体语义描述的资源推荐由于使用传统的推荐算法而存在用户冷启动、评分矩阵稀疏等问题导致推荐结果精度不高。本文提出一种基于本体推理规则和语义相似度计算的推荐方法，以解决推荐系统中存在的资源语义信息挖掘不充分、无法表达用户需求等问题，进而增强推荐系统的语义表达与处理能力。

2.2　本体规则推理

本体推理可以发现本体中隐含的逻辑关系、检查本体和知识的相容性、对实例进行自动分类[29]。这有利于保证本体构建的正确性和一致性，并可以将松散的概念、属性、实例等联系起来，形成一个完善的知识库，从而优化本体，降低本体维护的成本[30]。Hayes定义了D推导规则集，此规则集是从一个RDF图推导出一组标准推导规则[31]。由于D推导规则存在不完备性，Ter Horst对其进行补充[32]，称为D*推导规则集。这个规则集允许空白节点出现在三元组谓词位置上，以保证RDFS的完备性。OWL作为RDFS的扩展，解决了RDFS中存在的不支持基数约束、类的布尔组合、属性的限制等问题。Horrocks等将OWL1 DL与规则集成，基于Horn子句的RuleML中一个子集构造出语义网规则语言SWRL[33]。作为语义网框架，Jena也支持规则推理。Jena同SWRL类似，语法规则格式要求严格、与OWL描述方式相近，并且可以采用正向和反向推理，因此本文采用Jena规则对概念进行扩展推理。

2.3　语义相似度

语义相似度是指两个概念间的相似程度[17]，已经被应用于词义消歧[34]、自动检索[35-36]、图像分类及标注[37-38]、信息抽取[39]、信息检索[40-41]等领域。按照计算方法的不同分为：基于距离的方法、基于内容的方法和基于属性的方法等。基于距离的计算方法是在层次网络中使用路径长度来量化两个概念之间的语义距离[42]。两个概念的语义距离越大，即路径长度越长，则相似性越小。在层次网络中，全部有向边距离的权值都为1，即将各个节点视为同等重要，这样可以根据层次网络中构成最短路径的有向边数量计算两个概念的语义距离。该模型假设所有边权值都为1，但在实际情况中，节点的位置信息、节点的类型和节点之间的关联强度等因素都会影响其重要性，学者在此基础上对该模型进行改进。例如，Leacock等考虑了本体分类体系树自身的深度对概念相似度的影响，提出了改进的语义相似度计算模型[43]。基于内容的方法[44]认为两个概念共享的信息会影响二者的语义相似度。在层次网络中，概念子节点是对其父节点的细化和具体化，子节点包含父节点的信息内容，这样可以通过计算公共父节点概念所包含的信息内容计算子节点概念之间的相似度。基于属性的方法[45]是利用事物之间不同的属性特征区别事物。两个事物的公共属性越多，相似度越高。因此，可以利用两个概念对应的属性集的相似程度计算概念的语义相似度。然而，上述计算方法没有考虑本体结构信息，不能充分体现和揭示概念之间的语义关系，导致相似度计算的结构精度不高。本文提出的语义相似度计算方法，包含密度、深度和属性三种影响因素，可较为准确地计算出概念语义相似度。

3　推荐方法

3.1　推荐流程

基于本体规则推理和语义相似度计算的数字文献资源推荐方法是通过将用户输入概念扩展为一组相似概念，实现本体的细粒度查询，利用扩展后的概念与文献资源本身包含的语义信息进行语义相似度计算，将相似度较高的文献推荐给目标用户，从而向用户提供有价值的个性化资源推荐服务。构建的推荐流程如图1所示。首先用户输入关键词，利用本体知识库中的领域本体规则推理对输入关键词进行概念集扩展，得到扩展查询条件；其中资源层进行领域概念抽取，并根据抽取的领域知识创建本体，建立本体知识库；计算层将查询条件映射到本体实例(采用关键字匹配的方法)，计算扩展后概念集与资源的语义相似度；最

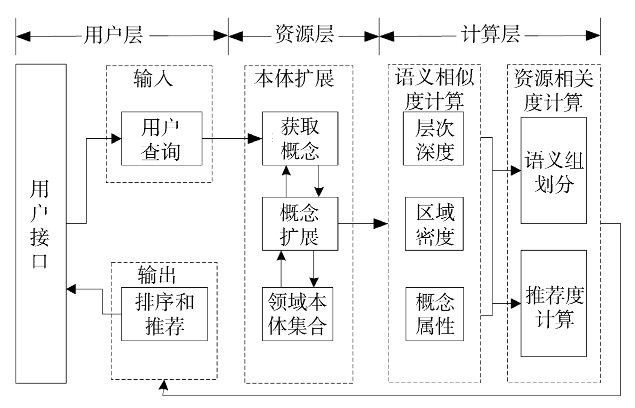


图 1　数字文献资源内容服务推荐流程

后，对推荐度进行排序，并将推荐度较高的数字文献推荐给目标用户。

3.2　算法设计及实现

(1) 语义推理

利用同义语义扩展进行语义推理，设定的规则包括概念上下位关系及相似关系。RDF作为语义网本体描述语言，规范了以三元组形式陈述的数据模型。RDFS可以用来表示简单的术语及其关系，例如类包含、属性包含、属性的定义域和值域等。由RDF和RDFS描述的推理规则具有反转性、传递性、继承性及部分性等。例如: {v p w|p rdfs:domain u.} =v rdfs:type u. (rdfs2)、{v p w|p rdfs:subPropertyOf q.} =v q w. (rdfs7)[20]。OWL作为RDFS的扩展，可以解决RDFS中存在的不支持基数约束、类的布尔组合、属性的限制等问题。OWL类和属性的推理规则具体包括：owl:sameAs、owl:intersectionOf、someValuesFrom 和allValuesFrom等。在OWL本体的定义中，利用owl:sameAs描述同义关系，由rdfs:subClassOf 描述上下位关系，并且二者都具有传递性。例如如果存在(?x rdfs: subClassOf ?y)，且(?y rdfs:subClassOf ?z)，那么则可以得到(?x rdfs: subClassOf ?z)。同样如果(?x owl:sameAs ?y)，且(?y owl:sameAs ?z)，则得到(?x owl:sameAs ?z)。虽然并未对x和z的关系进行在定义，但可以利用两个直接的定义推理得出二者隐含的定义，这就是推理机的作用。

Jena作为创建本体应用的Java框架结构，支持包括对RDF、RDFS、OWL等本体描述语言进行解析，对RDF文件和模型进行处理，对RDF模型持续性存储，基于规则的推理等功能。Jena提供基于规则的推理机(如RDFS Reasoner、OWL Reasoner等)包含传递推理、RDFS规则推理、OWL-Lite推理等推理功能，也包含通用规则推理和第三方推理引擎推理功能。Jena规则与规则推理机绑定，规则推理机通过调用bindSchema与模型或模式绑定[30]。由于Jena功能较为全面，因此选择Jena对本文设定的规则进行推理。

以下是对计算机推理的部分自定义产生式规则:

```
@prefixcomputer:<http://www.xh.com/computer.owl#>.
@include<RDFS>.
@include<OWL>.
String rules=
"[Rulel:(?x rdfs:subClassOf ?y), (?y rdfs:subClassOf ?z)->(?x
    rdfs:subClassOf ?z)]"+"[Rule2:(?x owl:sameAs ?y), (?y
    owl:sameAs ?z)->(?x owl:sameAs ?z)]"
//根据自定义推理规则创建对应的推理机
Reasoner reasoner=new GenericRuleReasoner
   (Rule.ParseRules(rules))
//根据自定义的推理机创建包含推理关系的数据模型
InfMfodel inf=ModelFactory.createlnfModel(reasoner, rawData)
```

(2) 语义相似度的计算

在领域本体构成的本体层次网络中，子节点是父节点概念的细化，子节点概念的含义比父节点更加具体。因此，概念所处位置深度越深，周围节点密度越大，表示概念包含的信息量越多。如果网络中子节点概念和父节点概念共同属性越多，那么二者关系相似度越高，有向边权重赋值越大。基于此，本文提出了语义相似度算法如公式(1)，其中影响因素包括层次深度、区域密度和概念属性。

$$sim(p_1,p_2)=\alpha\times\frac{D(Anc(p_1,p_2))\times\min(D(p_1),D(p_2))}{D(Anc(p_1,p_2))\times\min(D(p_1),D(p_2))+\omega\times len(p_1,p_2)}+\beta\times\frac{2\times IC[Anc(p_1,p_2)]}{IC(p_1)+IC(p_2)}+\gamma\times\frac{N(attr(p_1)\cap attr(p_2))}{N(attr(p_1)\cup attr(p_2))} \tag{1}$$

其中，$N(attr(p_1)\cap attr(p_2))$ 为p_1和p_2公共属性的数量。$\alpha+\beta+\gamma=1$，$\omega=\frac{N_{all}}{N_{anc}(p_1,p_2)}$，$N_{all}$为所有节点的数量，$N_{anc(p1,\ p2)}$为$p_1$和$p_2$共同祖先节点的数量。

基于本体的语义相似度改进算法的描述如下所示:

输入：抽象概念术语集合 ACS;
输出：待推荐文献资源及相似度;
Begin:
①For each ig∈抽象概念术语集合 ACS
②在 Glossary 表中查找与概念术语 ig 相似的概念存入 A1，并将数字文献资源概念放入字符串数组 A2
//Glossary 表为本文构建的本体中所有的概念集合

```
③While count(A1)≠0&&count(A2) ≠0
④从数组 A1 中取出一个概念和从数组 A2 中取出一个概念并
  计算其相似度 Simvalue
⑤If sim(ig, ig1)>阈值
⑥Do 扩展后抽象概念术语集合 EACSig=EACSig ∪ ig
⑦Simvalue=Simvalue+sim(ig, ig1)
   Endif
  Endwhile
End
```

(3) 推荐方法设计

本文提出的推荐方法可以用如下伪代码进行表示:

```
输入：用户查询概念 C;
输出：符合用户需求的数字文献资源;
Begin
①获得输入概念 c
②通过 Jena 进行语义推理，获得扩展的概念集合 C={c1, c2,…
  cn}
③查询资源库中包含 S(C)中的任意概念的文献资源，形成
  STMP(C, P)
④将扩展后的概念集合同描述文献资源的概念集合进行相关
  度计算
⑤按照公式 Recd=Simvalue/N 计算推荐度
  //N 为相似度的计算次数
⑥If Recd 大于阈值
⑦Do 用户设定推荐资源数量 r
    按照推荐度大小将 r 个资源 Items 推荐给用户
  Endif
End
```

4　实验设计及分析

4.1　本体构建

本文构建的计算机领域本体使用的概念术语主要来源于《计算机科学技术百科全书》[46]和《中国图书馆分类法》[47]。《计算机科学技术百科全书》收录了计算机领域概念中的完整术语，收词范围较广，词汇分类及定义规范，权威性强，因此将其作为构建计算机领域本体的参考。利用Protégé①中OntoGraf功能，可以显示构建本体的效果图。以计算机科学为中心形成可视的关联图如图2所示:

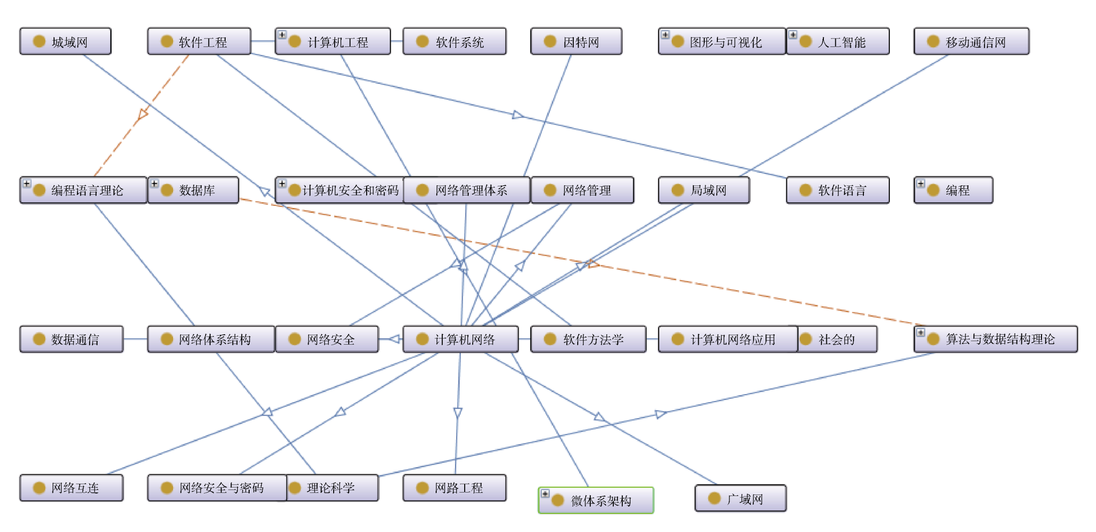


图 2　计算机领域本体部分示意图

本文构建的计算机领域本体包含1 000多个具有可检意义的计算机领域概念、属性及其相互关系。

4.2　算法实验

(1) 相似度检验

本文计算10组具有代表性概念的语义相似度，并与文献[42-43]提出的计算方法进行对比实验。另外，为了对比实验效果，采用咨询的方式获得人工对于语义相似度的判断。咨询对象包括计算机专业、情报专业、经济专业的硕士生和博士生，共有20人。通过对该组概念语义评价问题进行语义相似度判定。语义相似度的评判范围是[0-1], 0表示两个概念完全不同, 1表示两个概念语义相同。对专业及非专业的受测者各进行两次试验，并对同一概念语义相似度的评测结果取平均值。

表1中的Sim1和Sim2为文献[42-43]语义相似度方法计算得到的结果；Sim3为本文提出的方法进行

①http://protege.stanford.edu/.

语义相似度计算后得到的结果，最后一项为人工判断的结果。

表 1　概念语义相似度计算结果

词汇 1	词汇 2	Sim1	Sim2	Sim3	人工判断
信息	信念	0.5	0.416	0.468	0.465
软件	C 语言	0.333	0.380	0.362	0.357
软件	面向对象	0.417	0.385	0.492	0.507
软件	构件	0.571	0.568	0.435	0.422
电路	存储器	0.736	0.752	0.603	0.621
加法器	寄存器	0.375	0.522	0.303	0.299
网络通信	光纤	1.000	1.000	0.804	0.778
局域网	以太网	0.231	0.230	0.205	0.109
图形学	等值面	0.387	0.397	0.285	0.291
多媒体	流媒体	0.380	0.388	0.355	0.357

由表1可知，对比分析Sim1、Sim2、Sim3这三列，前5行可以看出这三种算法计算结果均比较符合目标用户的主观判断，但是对于概念节点附近密度较大或者公共属性较多的概念，会导致相似度计算结果并不合理，而本文提出的方法能够较为准确地计算出概念相似度，使语义相似度计算结果符合目标用户的主观判断。

(2) 推荐检验

为了检验该推荐方法，从知网下载800篇计算机领域的文献作为数据集，这些文献包含题目、摘要和关键词。由于关键词可以描述文章的主题，因此选用关键词作为文章的语义描述。利用F评价指标比较不同数据集的推荐效果，F评价指标[48]包括查准率(Precision)和查全率(Recall)两项指标，F值越高，推荐效果越好。

$$F = 2 \times Precision \times Recall / (Precision + Recall) \qquad (2)$$

其中, Precision 与 Recall 计算方法分别为:

$$Precision = \frac{\text{符合用户需求的推荐的数字文献资源数量}}{\text{推荐的数字文献资源总数}} \times 100\% \qquad (3)$$

$$Recall = \frac{\text{推荐的数字文献资源数量}}{\text{数字文献资源总数}} \times 100\% \qquad (4)$$

实验结果如表 2 所示。由表 2 可知，采用本文方法得到的文献资源推荐结果 F 值随着文献数量的增加而提高。由于充分考虑了本体知识库中关于概念密度和概念属性等本体的结构特征，使得本文的概念相似度计算方法计算得到的结果可以更准确地反映文献资源间的语义相似性，丰富了文献资源语义信息，提高了语义相关度计算效果，并对目标用户进行有效的推荐。

表 2　本文算法 F 评价指标

测试集数据	测试数字文献资源	数字文献资源推荐总数	成功推荐数目	P	R	F 值
10%	80	40	16	0.400	0.5	0.44
30%	240	120	54	0.450	0.5	0.477
50%	400	200	93	0.465	0.5	0.482
70%	560	280	142	0.507	0.5	0.504
90%	720	360	201	0.558	0.5	0.528

5　结　语

本文提出了一种基于本体规则推理和语义相似度计算的数字文献资源内容服务推荐方法。该方法利用 Jena 规则推理对用户输入的概念进行语义扩展，将扩展后的概念集与数字文献资源自身语义信息进行相似度计算和排序，得出推荐度较高的文献并推荐给目标用户。结果表明，该方法能够比较准确地反映概念之间的语义关系及计算概念节点之间的语义相似度，可以充分挖掘用户需求并形成具有针对性的数字资源内容服务推荐。由于采集的文献资源数量和内容丰富程度不够，并且推荐结果由用户判断其推荐的准确程度，具有一定的主观性；为了对比推荐效果，本文的查全率设定为 0.5，并没有随着测试文献的数量增加而增加，以上导致了本文 F 值低于现有的推荐算法。因此，提高推荐方法的 F 值及推荐文献资源的丰富程度也是今后研究的重点。

参考文献：

[1] 马炎. 一种自适应的协作过滤图书推荐系统研究[J]. 情报杂志，2008，27(5)：105-109. (Ma Yan. Research on the Adaptive Collaborative Filtering Recommendation System [J]. Journal of Information, 2008, 27(5): 105-109.)

[2] 董坤. 基于协同过滤算法的高校图书馆图书推荐系统研究[J]. 现代图书情报技术，2011(11)：44-47. (Dong Kun. Research of Personalized Book Recommender System of University Library Based on Collaborative Filter [J]. New Technology of Library and Information Service, 2011(11): 44-47.)

[3] 吴志强，马慧娟. 协同信息推荐技术及其在数字图书馆中的应用研究述评[J]. 图书情报工作，2012，55(19)：122-127.

(Wu Zhiqiang, Ma Huijuan. Review on Researches About the Application of Collaborative Information Recommendation Technologies in Digital Libraries [J]. Library and Information Service, 2012, 55(19): 122-127.)

[4] 熊拥军. 数字图书馆个性化服务资源推荐模式分析[J]. 图书馆, 2014(2): 132-134. (Xiong Yongjun. The Model Analysis of Personalized Information Recommendation Service in Digital Library [J]. Library, 2014(2): 132-134.)

[5] Koren Y, Bell R, Volinsky C. Matrix Factorization Techniques for Recommender Systems [J]. Computer, 2009, 42(8): 30-37.

[6] Ghazarian S, Nematbakhsh M A. Enhancing Memory-based Collaborative Filtering for Group Recommender Systems[J]. Expert Systems with Applications, 2015, 42(7): 3801-3812.

[7] Zhu Z, Wang J Y. Book Recommendation Service by Improved Association Rule Mining Algorithm [C]. In: Proceedings of the 6th International Conference on Machine Learning and Cybernetics. 2007: 19-22.

[8] Giseli Rabello Lopes, Maria Aparecida Martins Souto, Leandro Krug Wives, et al. A Personalized Recommender System for Digital Libraries [C]. In: Proceedings of the 14th Brazilian Symposium on Multimedia and the Web. 2008: 59-66.

[9] Amini B, Ibrahim R, Othman M S, et al. Incorporating Scholar's Background Knowledge into Recommender System for Digital Libraries [C]. In: Proceedings of the 5th Malaysian Conference in Software Engineering (MySEC). 2011: 516-523.

[10] Will T C, Srinivasan A, Wu Y F B. Search Personalization: Knowledge-Based Recommendation in Digital Libraries [C]. In: Proceedings of Americas Conference on Information Systems. 2009: 728-735.

[11] IJntema W, Goossen F. Ontology-Based News Recommendation [C]. In: Proceedings of the 1st International Workshop on Data Semantics, Switzerland. 2010: 22-26.

[12] 曾春, 邢春晓, 周立柱. 个性化服务技术综述[J]. 软件学报, 2002, 13(10): 1592-1601. (Zeng Chun, Xing Chunxiao, Zhou Lizhu. A Survey of Personalization Technology [J]. Journal of Software, 2002, 13(10): 1592-1601.)

[13] Chen Y J, Chu H C, Chen Y M, et al. Adapting Domain Ontology for Personalized Knowledge Search and Recommendation [J]. Information & Management, 2013, 50(6): 285-303.

[14] Wu H, Yue K, Pei Y, et al. Collaborative Topic Regression with Social Trust Ensemble for Recommendation in Social Media Systems [J]. Knowledge-Based Systems, 2016, 97(1): 111-122.

[15] Isinkaye F O, Folajimi Y O, Ojokoh B A. Recommendation Systems: Principles, Methods and Evaluation [J]. Egyptian Informatics Journal, 2015, 16(3): 261-273.

[16] 吴正洋, 汤庸, 方家轩, 等. 一种基于本体语义相似度的协同过滤推荐方法[J]. 计算机科学, 2015, 42(9): 204-207, 225. (Wu Zhengyang, Tang Yong, Fang Jiaxuan, et al. Collaborative Filtering Recommendation Algorithm Based on Ontology Semantic Similarity [J]. Computer Science, 2015, 42(9): 204-207, 225.)

[17] 刘宏哲, 须德. 基于本体的语义相似度和相关度计算研究综述[J]. 计算机科学, 2012, 39(2): 8-13. (Liu Hongzhe, Xu De. Ontology Based Semantic Similarity and Relatedness Measures Review [J]. Computer Science, 2012, 39(2): 8-13.)

[18] 何超, 张玉峰. 基于本体的馆藏数字资源语义聚合与可视化研究[J]. 情报理论与实践, 2013, 36(10): 73-76. (He Chao, Zhang Yufeng. Research on Semantic Aggregation and Visualization of Library Digital Resources Based on Ontology [J]. Information Studies: Theory & Application, 2013, 36(10): 73-76.)

[19] 乔冬春, 刘晓燕, 付晓东, 等. 一种基于本体的推荐系统模型[J]. 计算机工程, 2014, 40(11): 282-287. (Qiao Dongchun, Liu Xiaoyan, Fu Xiaodong, et al. An Ontology-based Recommendation System Model [J]. Computer Engineering, 2014, 40(11): 282-287.)

[20] Chen R C, Huang Y H, Bau C T, et al. A Recommendation System Based on Domain Ontology and SWRL for Anti-diabetic Drugs Selection [J]. Expert Systems with Applications, 2012, 39 (4): 3995-4006.

[21] Vesin B, Ivanović M, Klašnja-Milićević A, et al. Protus 2. 0: Ontology-based Semantic Recommendation in Programming Tutoring System [J]. Expert Systems with Applications, 2012, 39(15): 12229-12246.

[22] Torshizi A D, Zarandi M H F, Torshizi G D, et al. A Hybrid Fuzzy-ontology Based Intelligent System to Determine Level of Severity and Treatment Recommendation for Benign Prostatic Hyperplasia [J]. Computer Methods and Programs in Biomedicine, 2014, 113(1): 301-313.

[23] Hsu I C. Integrating Ontology Technology with Folksonomies for Personalized Social Tag Recommendation [J]. Applied Soft Computing, 2013, 13(8): 3745-3750.

[24] 肖敏. 基于领域本体的电子商务推荐技术研究[D]. 武汉: 武汉理工大学, 2009. (Xiao Min. Research on Electronic Commerence Recommendation Technology Base on Domin Ontology [D]. Wuhan: Wuhan University of Technology,

2009.)

[25] 李言则. 基于课程本体的学习内容个性化推荐的研究与应用[D]. 武汉: 华中师范大学, 2013. (Li Yanze. Research and Application of Personalized Leaning Content Recommendation Based on Course Ontology [D]. Wuhan: Central China Normal University, 2013.)

[26] Al-Nazer A, Helmy T, Al-Mulhem M. User's Profile Ontology-based Semantic Framework for Personalized Food and Nutrition Recommendation [J]. Procedia Computer Science, 2014, 32: 101-108.

[27] Yu Y H, Kim J H, Shin K, et al. Recommendation System Using Location-based Ontology on Wireless Internet: An Example of Collective Intelligence by Using 'Mashup' Applications [J]. Expert Systems with Applications, 2009, 36(9): 11675-11681.

[28] Weng S S, Chang H L. Using Ontology Network Analysis for Research Document Recommendation [J]. Expert Systems with Applications, 2008, 34(3): 1857-1869.

[29] 龚资. 基于 OWL 描述的本体推理研究[D]. 长春: 吉林大学, 2007. (Gong Zi. Research on Ontology Reasoning Based on OWL [D]. Changchun: Jilin University, 2007.)

[30] Hebeler J, Fisher M, Blace R, et al. Semantic Web Programming [M]. Wiley Publishing, Inc. 2009.

[31] Hayes P. RDF Semantics [OL]. [2016-04-04]. http://www.w3.org/TR/rdf-mt/.

[32] Ter Horst H J. Completeness, Decidability and Complexity of Entailment for RDF Schema and a Semantic Extension Involving the OWL Vocabulary [J]. Web Semantics: Science, Services and Agents on the World Wide Web, 2005, 3(2-3): 79-115.

[33] Horrock I, Patel-Schneider P F. Reducing OWL Entailment to Description Logic Satisfiability [C]. In: Proceedings of the International Semantic Web Conference. 2005: 345-357.

[34] Singh S, Siddiqui T J. Role of Semantic Relations in Hindi Word Sense Disambiguation [J]. Procedia Computer Science, 2015, 46: 240-248.

[35] Alonso I, Contreras D. Evaluation of Semantic Similarity Metrics Applied to the Automatic Retrieval of Medical Documents: An UMLS Approach [J]. Expert Systems with Applications, 2016, 44 (C): 386-399.

[36] Hussain S F, Suryani A. On Retrieving Intelligently Plagiarized Documents Using Semantic Similarity [J]. Engineering Applications of Artificial Intelligence, 2015, 45: 246-258.

[37] Chang J Y, Lee K M. Large Margin Learning of Hierarchical Semantic Similarity for Image Classification [J]. Computer Vision and Image Understanding, 2015, 132: 3-11.

[38] Zhang X, Liu C. Image Annotation Based on Feature Fusion and Semantic Similarity [J]. Neurocomputing, 2015, 149: 1658-1671.

[39] Hassan H, Hassan A, Emam O. Unsupervised Information Extraction Approach Using Graph Mutual Reinforcement[C]. In: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. 2006: 501-508.

[40] Kalloubi F, Nfaoui E H, Beqqali O E. Microblog Semantic Context Retrieval System Based on Linked Open Data and Graph-based Theory[J]. Expert Systems with Applications, 2016, 53: 138-148.

[41] Bae M, Kang S, Oh S. Semantic Similarity Method for Keyword Query System on RDF[J]. Neurocomputing, 2014, 146: 264-275.

[42] Rada R, Mili H, Bicknell E, et al. Development and Application of a Metric on Semantic Nets [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1989, 19(1): 17-30.

[43] Leacock C, Chodorov M. Combining Local Context and WordNet Similarity for Word Sense Identification [A]. // Fellbaum C, Miller G. WordNet: An Electronic Lexical Database and Some of Its Applications [M]. MIT Press, 1998: 265-283.

[44] Lord P W, Stevens R D, Brass A, et al. Investigating Semantic Similarity Measures Across the Gene Ontology: The Relationship Between Sequence and Annotation [J]. Bioinformatics, 2003, 19(10): 1275-1283.

[45] Tversky A. Feature of Similarity [J]. Psychological Review, 1977, 84(4): 327-352.

[46] 张效祥. 计算机科学技术百科全书[M]. 北京: 清华大学出版社, 2005. (Zhang Xiaoxiang. Encyclopedia of computer science and technology [M]. Beijing: Tsinghua University Press, 2005.)

[47] 国家图书馆《中国图书馆分类书》编辑委员会. 中国图书馆分类法[M]. 北京: 国家图书馆出版社, 2010. (Editorial Board of《Chinese Library Classification》of National Library. Chinese Library Classification [M]. Beijing: National Library of China Publishing House, 2010.)

[48] Zhou T, Kuscsik Z, Liu J G, et al. Solving the Apparent Diversity-accuracy Dilemma of Recommender Systems [J]. Proceedings of the National Academy of Sciences, 2010, 107(10): 4511-4517.

作者贡献声明:

毕强: 提出研究命题及研究思路;
刘健: 论文撰写, 数据处理及实证研究;
刘庆旭: 英文摘要撰写及修改;
王福: 摘要撰写及论文修改。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail: tomosliu9999@126.com。
[1] 刘健, 毕强. Glossary.csv. 本体概念集合.
[2] 刘健, 毕强. CNKI.resoures.sql. 知网下载资源数据库文件.
[3] 刘健, 毕强. Similarity.doc. 相似度调查表.

收稿日期: 2016-05-09
收修改稿日期: 2016-06-19

New Content Recommendation Service of Digital Literature

Liu Jian[1] Bi Qiang[1] Liu Qingxu[1] Wang Fu[1, 2]
[1](School of Management, Jilin University, Changchun 130022, China)
[2](Inner Mongolia University of Technology Library, Huhhot 010051, China)

Abstract: [**Objective**] This paper tries to improve the traditional content recommendation service of digital literature, which cannot fully exploit the semantic information of the literature. [**Methods**] First, we introduced the Ontology reasoning rules to the recommendation system, and then semantically extended the user's query. Second, we calculated the similarity of the literature to rank. Finally, we recommend those top ranked literature to the users. [**Results**] The proposed algorithm can calculate the semantic similarity among literature and successful recommend documents to the users. [**Limitations**] Only examined the new method with relatively small data sets. [**Conclusions**] The proposed algorithm could effectively exploit the semantic information of target literature and offer a new way to recommend digital resource to the users.

Keywords: Digital literature Service recommendation Ontology reasoning Semantic similarity